Risk Analysis Overview

Including "*Bad Things Good Applicants Do*"

Federal Aviation Administration

Presented to: 2006 Dallas DER Seminar

By: Ann Azevedo, CSTA, Aircraft Safety Analysis

Date: May 25, 2006

Topics

- Risk analysis principles – brief overview
- Bad things good applicants do
- Discussion or questions

Contact Information

Ann Azevedo

- ANE-104
  12 New England Executive Park
  Burlington, MA 01803
- 781-238-7117
- Ann.Azevedo@faa.gov (best way)

Assessing Risk

- Like any other discipline, risk assessment is a specialized process
- However, like any other discipline, all should have a basic understanding of techniques

FAA Philosophy

We should aim to respond to threats to continued operational safety with a timeliness and manner in keeping with the severity and probability of the event

Assessing Risk

Components Of Risk

- Severity
- Probability

Severity

- Basic event is not necessarily the event of interest
- Conditional probability of more severe event

Questions Risk Analysis Will Answer

- Severity
  - What could happen?
  - What are the potential outcomes?
- Probability
  - How often will it happen?
  - What are the probabilities of those outcomes?
- If we are speaking of risk *management*, we need one more question answered:
  - What is acceptable?

The Benefits of Risk Analysis

- Ensures that continued operation is acceptable
- Allows for objective and consistent assessment of unsafe conditions
- Allows for mitigation of unsafe conditions with the most optimum use of resources within an individual problem
- Helps to prioritize use of resources among multiple problems

Assessing Risk – Failure Distributions

Random

- Failures equally likely whatever the age of the component (Ex. - fan blade fractures due to bird strikes)
- Many failure modes combined often result in a random distribution (Ex. - IFSDs)
- Future risk easy to calculate:
  - Mean Time Between Failure (MTBF) from past data (total hours or cycles divided by number of events)
  - Divide future hours/cycles by MTBF for number of expected future events

Failure Distributions (Cont.)

Wearout

- Failures become more likely the older the component gets (Ex. - low cycle fatigue)
  - Rate of increasing probability of failure can vary
- Most common failure mode for hardware
- Future risk more complicated to calculate – function of current age on each individual part
  - If Weibull distribution, probability of failing within next x hours/cycles (given current age t) = [P(t+x) - P(t)] / [1 - P(t)]

Failure Distributions (Cont.)

Infant Mortality

- Failures become less likely the older the component gets (Ex. - Maintenance, assembly errors)
  - Older parts may become exempt from suspicion
- Often difficult to manage risk
- Future risk function of current age on each individual part

Failure Distributions (Cont.)

Other failure distributions of interest

- Binomial
  - Ex. - Modeling presence of latent failures
  - Yes/No failure does *not* (necessarily) mean 50-50 probability!
- Poisson
  - Models events
  - The probability of actually having an event given a risk factor (future number of events) is a Poisson function (Ex. - a 0.7 risk factor equals a 50% probability of at least 1 event)

Managing Risk

- We *manage* risk – this implies we have something to manage it *to* – that is, we need guidelines for what is acceptable risk
- Zero risk is unattainable without immediate intervention (grounding)

Risk Analysis Objectives

Provides a systematic means to

- **Identify**
- **Assess**
- **Prioritize**

safety threats

Bad Things Good Applicants Do

In Certification

- ‘Identicality’ demonstrations
- Fault trees – blanket assumption of independence of failures
- 10^{-9} for single failures

In Continued Airworthiness

- Worst-case assessments instead of total risk
- Assuming randomness
- Treating failures as ‘outliers’

Other Issues

In Continued Airworthiness

- “I have no data”
- Rates vs. possible events
- Adding probabilities of failure
- Confidence and statistical significance

Identicality Demonstrations

- Applicants present two populations – theirs and that to which they must show ‘identicality’ through test and computation
- Applicants typically perform statistical test that is designed to test for *differences* – not for identicality!
- In other words, the test they use is to “prove that it’s different”, not to “show that it’s the same” – this is a very different question, and requires a much larger sample size
- Work-around – require that parts by the new applicant fall within the range of actual measurements from the other sample

Fault Trees and Independence

- Applicants multiply failure probabilities together
- This assumes independence of those failures
- Has the rationale for this assumption been presented?

Note: Assumption of independence is not necessarily incorrect (in fact, it's probably usually correct, but not always)

Single Failures and 10^{-9}

Very small numbers cannot be demonstrated ahead of time for single failures

- Can't test to 10^{-9}
- Parts typically are not in common usage, so little prior experience (or specifications) exist
- If estimate is wrong, the first warning you get is a serious incident (if a hazardous or catastrophic failure)

Prior service history with similar parts provides some assurance, but is not a guarantee

Must rely on best practices

Worst Case

- Assessment of risk presented based on the ‘worst case’ component
- Assumption that, if the ‘worst case’ has acceptable risk, all others are okay, too
- However, we are concerned about the *total* risk. While the ‘worst case’ might be the biggest single part of the risk, the total risk from all the components, added together, may be unacceptable
- This can sometimes be a certification issue, as well

Blanket Assumption of Random Failure Rate

- As we discussed earlier, random failure rates are easy to deal with
- As a result (or because of presumed lack of data), applicants often use the random failure rate statistics to estimate risk
- However, this can result in a serious under-estimation of risk
- Workaround – if there's nothing else to use, only go back 1 or 2 years and only predict 1 or 2 years into the future

It’s an Outlier!

- Failures that are not understood (never seen before) are treated as one-offs (assume will not repeat)
- Tendency to use extenuating circumstances (operational, material, etc.) to eliminate failure(s) from the data under discussion

It’s an Outlier! (Cont.)

- The applicant will want to eliminate failure(s) from the numerator, but will not reduce the successful operations with the extenuating circumstances from the denominator
  - For example, applicant says “this failure doesn’t count because the airplane was operated in this condition (or location),” but doesn’t then discount hours by all the successful operation in that condition

“I Have No Data”

Therefore, I refuse to do an analysis?

- Workarounds:
  - Production shipping records
  - Aircraft sales information
  - Utilization data from information on the failures

“But I’ve Only Had 1 Failure, And They Had 7!”

- Actual numbers of failures (accidents, etc.) do not mean much without knowing the size and experience of the full population
- Which is worse?
  - 7 failures in a fleet of 3,000 airplanes that have been in service for an average of 10 years?
  - 1 failure in a fleet of 50 airplanes that have been in service for 2 years?
- Use *rates* of occurrence (preferably, compare failure distributions), not actual numbers

Adding Probabilities

Failure A probability + failure B probability, right?

P(A) of failure ‘A’ occurring

P(B) of failure ‘B’ occurring

What’s the probability that I get a failure?

How many say P(A) + P(B)?

Adding Probabilities (cont.)

Suppose:

P(A) = 40% (.40)

P(B) = 75% (.75)

So P(A) + P(B) = .40 + .75 = 1.15

115% probability of failure???

Adding Probabilities (cont.)

Actual formula for summing failures is:

$$1 – (1 – P(A))(1 – P(B))$$

Our example is:

$$1 – (1 – 0.40)(1 – 0.75) = 0.85$$

So why do people add probabilities?

Adding Probabilities (cont.)

- For small numbers (< .05), adding the probabilities is a good approximation for the true formula
  - Example: P(A) = 0.01, P(B) = 0.02
  - True formula: $P = 1 - (1 - 0.01)(1 - 0.02) = .0298$
  - Approximation: $P = 0.01 + 0.02 = .03$
- Problem arises because people use the approximation (thinking it is the correct formula) for all probabilities

Confidence

‘Confidence bands’ are a statistical indication of the percent of the time the *actual* answer (average or other parameter of interest) falls within a specified range of the *estimated* answer

Confidence (Cont.)

- 95% confidence bands – 95% of the time (that is, in 95 out of 100 similar cases), the actual answer will lie within the specified bounds around the estimated answer
- More likely that the actual answer lies closer to the estimated answer rather than closer to the edge of the confidence band
- Confidence bands vary inversely with the size of the sample – a larger sample will give a better estimate of the actual answer

Confidence (cont.)

We can also have 'confidence' that an observed difference between two populations is a real difference and not merely a chance result when no difference really exists

- 95% confident that there is a difference – 95% of the time, there is a true difference between the two populations
- The larger the sample size, the smaller the difference that we can detect as being 'statistically significant' (all else being equal)

Statistically Significant

- 'Statistical significance' means that a difference exists between the two populations with a given degree of confidence (usually 95% or 99%)
- Things can be *statistically* significant but not *practically* significant
  - For example, we may only want to switch to a new (and more expensive) process if it improves yield by more than 5%. We may find a statistically significant difference of 2%, but that will not have practical significance for us

Warning!

- These statements on confidence and significance are all dependent on the assumption that the sample you have is *representative*
  - Not just one heat or lot or operator (ignores differences between those groups)
- Methodologies to get a representative sample
  - *Random* – sample selected without exclusion of any particular component, and without bias
  - *Stratified* – sample adjusted to ensure that different components are selected in a percentage that matches their true breakdown

Discussion or Questions?